

Distinguishing noise from chaos: objective versus subjective criteria using Horizontal Visibility Graph.

Martín Gómez Ravetti ¹, Laura C. Carpi ², Bruna Amin Gonçalves¹, Alejandro C. Frery ², Osvaldo A. Rosso ³

1 Departamento de Engenharia de Produção, Universidade Federal de Minas Gerais.
Av. Antônio Carlos, 6627, Belo Horizonte, 31270-901 Belo Horizonte - MG, Brazil.

2 LaCCAN/CPMAT - Instituto de Computação, Universidade Federal de Alagoas
BR 104 Norte km 97, 57072-970 Maceió, Alagoas - Brazil.

3 Instituto de Física, Universidade Federal de Alagoas.
BR 104 Norte km 97, 57072-970 Maceió, Alagoas - Brazil.
Laboratorio de Sistemas Complejos, Facultad de Ingeniería,
Universidad de Buenos Aires (UBA).

(1063) Av. Paseo Colón 840, Ciudad Autónoma de Buenos Aires, Argentina.

* E-mail: martin.ravetti@dep.ufmg.br

Abstract

A recently proposed methodology called the Horizontal Visibility Graph (HVG) [Luque *et al.*, Phys. Rev. E., 80, 046103 (2009)] that constitutes a geometrical simplification of the well known Visibility Graph algorithm [Lacasa *et al.*, Proc. Natl. Sci. U.S.A. 105, 4972 (2008)], has been used to study the distinction between deterministic and stochastic components in time series [L. Lacasa and R. Toral, Phys. Rev. E., 82, 036120 (2010)]. Specifically, the authors propose that the node degree distribution of these processes follows an exponential functional of the form $P(\kappa) \sim \exp(-\lambda \kappa)$, in which κ is the node degree and λ is a positive parameter able to distinguish between deterministic (chaotic) and stochastic (uncorrelated and correlated) dynamics. In this work, we investigate the characteristics of the node degree distributions constructed by using HVG, for time series corresponding to 28 chaotic maps and 3 different stochastic processes. We thoroughly study the methodology proposed by Lacasa and Toral finding several cases for which their hypothesis is not valid. We propose a methodology that uses the HVG together with Information Theory quantifiers. An extensive and careful analysis of the node degree distributions obtained by applying HVG allow us to conclude that the Fisher-Shannon information plane is a remarkable tool able to graphically represent the different nature, deterministic or stochastic, of the systems under study.

1 Introduction

The distinction between stochastic and chaotic processes has received much attention, becoming one of the most appealing problems in time series analysis. Since stochastic and chaotic (low dimensional deterministic) processes share several characteristics, the discrimination between them is a challenging task. Time series with complex structures are very frequent in both natural and artificial systems. The interest behind this distinction relies in uncovering the cause of unpredictability governing these systems.

Much effort has been dedicated in the understanding of this topic. It was thought, in the origins of chaotic dynamics, that obtaining finite, non-integer values for fractal dimension was a strong evidence of the presence of deterministic chaos, as stochastic processes were thought to have an infinite value. Osborne and Provenzale [1] observed for a stochastic process a non-convergence in the correlation dimension (as a estimation of fractal dimension). They showed that time series generated by inverting power law spectra and random phases are random fractal paths with finite Hausdorff dimension and, consequently, with finite correlation dimension [1].

Among other methodologies to distinguish chaotic from stochastic time series we can mention the work of Sugihara and May [2] based on nonlinear forecasting in which they compare predicted and actual trajectories and make tentative distinctions between dynamical chaos and measurement errors. The accuracy of nonlinear forecast diminishes for increasing prediction time-intervals for a chaotic time series. This dependency is not found for uncorrelated noises [2]. Kaplan and Glass [3, 4] observed that the tangent to the trajectory generated by a deterministic system is a function of the position in phase space, consequently, all the tangents to a trajectory in a given phase space region will display similar orientation. As stochastic dynamics do not exhibit this behavior, Kaplan and Glass proposed a test based on these observations. Kantz and co-workers [5, 6] recently analyzed the behavior of entropy quantifiers as a function of the coarse-graining resolution, and applied their ideas to distinguish between chaos and noise. Their methodology can be considered a generalization of the Grassberger and Procaccia method [7] regarding the estimation of the correlation dimension and the consideration of finite values as signatures of deterministic behavior.

Chaotic systems display “sensitivity to initial conditions” which manifests instability everywhere in the phase space and leads to non-periodic motion (chaotic time series). They display long-term unpredictability despite the deterministic character of the temporal trajectory. In a system undergoing chaotic motion, two neighboring points in the phase space move away exponentially rapidly. Let $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t)$ be two such points, located within a ball of radius R at time t . Further, assume that these two points cannot be resolved within the ball due to poor instrumental resolution. At some later time t' the distance between the points will typically grow to $|\mathbf{x}_1(t') - \mathbf{x}_2(t')| \approx |\mathbf{x}_1(t) - \mathbf{x}_2(t)| \exp(\Lambda |t' - t|)$, with $\Lambda > 0$ for a chaotic dynamics, being Λ the biggest Lyapunov exponent. When this distance at time t' exceeds R , the points become experimentally distinguishable. This implies that instability reveals some information about the phase space population that was not available at earlier times [8]. The above considerations allow to think chaos as an *information source*. Moreover, the associated rate of generated information can be formulated in a precise way in terms of Kolmogorov-Sinai’s entropy [9, 10].

In more recent works, the use of quantifiers based on Information Theory has led to interesting results regarding the characteristics of nonlinear chaotic dynamics, improving the understanding of their associated time series. Specifically, Rosso and collaborators [11–14] found that the use of the normalized Shannon entropy and statistical complexity allow for a better distinction between stochastic and chaotic dynamics, when incorporating causal information via the Bandt and Pompe methodology [15, 16], generating in this way the graphic tool called *the causality entropy-complexity plane*. In the same fashion, Olivares *et al.* [17, 18] propose the use of two information quantifiers like measures, namely, the normalized Shannon entropy and the Fisher information combined in the so-called *the causality Shannon-Fisher plane*, finding that the two dynamics exhibit different planar locations.

The use of Visibility Graphs (VG) introduced by Lacasa and co-workers [19], a method that transform a time series into a graph, has also been used with this purpose. Specifically, HVGs, a geometrical simplification of VGs, which is also computationally faster, was applied in the classification and characterization of periodic, chaotic, and onset of chaos dynamics [20, 21].

Lacasa and Toral [22] studied the discrimination between chaotic, uncorrelated and correlated stochastic time series by using HVG. They conjecture that the node degree distribution of these systems follows an exponential functional of the form $P(\kappa) \sim \exp(-\lambda \kappa)$, in which λ is a positive parameter and κ the node degree. They computed analytically the HVG-PDF for the case of uncorrelated noise (white noise) [23], and found the corresponding parameter value $\lambda_c = \ln(3/2)$. Moreover, they hypothesized that this value corresponds to a central value that separates correlated stochastic ($\lambda > \lambda_c$) from chaotic dynamics ($\lambda < \lambda_c$).

Even though the methodology works for several chaotic and stochastic systems, we have found several examples for which results diverge from the ones expected.

In this work, we present a methodology able to discriminate between chaotic and stochastic (uncorrelated and correlated) time series by using the HVG methodology together with Information Theory

quantifiers. A total of 31 systems are considered; the 27 chaotic maps described by Sprott [24] and the Schuster map [25], noises with f^{-k} , $k \geq 0$ power spectrum (PS) and stochastic time series generated by fractional Brownian motion (fBm) and fractional Gaussian noise (fGn).

Following Olivares *et al.* [17, 18] we based our analysis on the so-called Shannon-Fisher information plane ($\mathcal{S} \times \mathcal{F}$) that captures both global and local features of the system's dynamics. Its horizontal and vertical axis are functionals of the pertinent probability distribution, namely, the normalized Shannon entropy (\mathcal{S}) and the normalized Fisher Information measure (\mathcal{F}). We evaluate these quantifiers for the time series using as PDF the node degree distribution obtained via the horizontal visibility graph. We show that the Shannon-Fisher information plane is able to efficiently represent the different nature of the systems in a planar representation, as well as to distinguish between the different degrees of correlation structures.

As for the organization of this work, the forthcoming Section 2.1 enumerates and describes the chaotic maps and the stochastic processes considered. Section 2.2 describes the Horizontal Visibility Graph algorithm and discusses the characterization of the HVG-PDF from a statistical point of view, as well as the methodology implemented in [22] based on the parameter λ . In Section 2.3 the basis on the Shannon-Fisher plane is detailed, and finally, Section 3 present our results and discussions.

2 Materials and Methods

2.1 Chaotic maps and stochastic processes

2.1.1 Chaotic maps

In the present work we consider 27 chaotic maps described by Sprott in his book [24] and the Schuster Maps [25] (see supplementary material), grouped as follows:

noninvertible maps: (1) Logistic map [26]; (2) Sine map [27]; (3) Tent map [28]; (4) Linear congruential generator [29]; (5) Cubic map [30]; (6) Ricker's population model [31]; (7) Gauss map [32]; (8) Cusp map [33]; (9) Pinchers map [34]; (10) Spence map [35]; (11) Sine-circle map [36].

dissipative maps: (12) Hénon map [37]; (13) Lozi map [38]; (14) Delayed logistic map [39]; (15) Tinkerbell map [40]; (16) Burgers' map [41]; (17) Holmes cubic map [42]; (18) Dissipative standard map [43]; (19) Ikeda map [44]; (20) Sinai map [45]; (21) Discrete predator-prey map [46].

conservative maps: (22) Chirikov standard map [47]; (23) Hénon area-preserving quadratic map [48]; (24) Arnold's cat map [49]; (25) Gingerbreadman map [50]; (26) Chaotic web map [51]; (27) Lorenz three-dimensional chaotic map [52].

We also analyze the Schuster map, a class introduced by Schuster and co-workers [25] that exhibits intermittent signals with chaotic bursts and f^{-z} power spectrum (PS). It is defined as:

$$x_{n+1} = (x_n + x_n^z) \bmod 1. \quad (1)$$

As the parameter z increases, the laminar zone increases in size and the chaotic bursts are less frequent. To generate these maps we use random initial conditions in the interval $(0, 1)$ and we consider $z \in \{1.25, 1.50, 1.75, 2.00\}$.

For noninvertible, dissipative and conservative maps we use the initial conditions and parameter-values detailed by Sprott. The corresponding initial values are given in the basin of attraction for noninvertible maps, near the attractor for dissipative maps, and in the chaotic sea for conservative maps [24]. In the generation process N iterations were considered after discarding the first 10^5 . In the case of multi-dimensional maps, we consider all map coordinates. A complete description of these maps can be found in the supplementary material.

2.1.2 Stochastic processes

The following classical stochastic processes are considered in this work:

Noises with f^{-k} power spectrum generated as follows: 1) By using the Mersenne twister generator [54], the MATLAB[®] *RAND* function is used to produce pseudo random numbers in the interval $(-0.5, 0.5)$ with an almost flat power spectrum (PS), uniform PDF, and zero mean value.

2) The Fast Fourier Transform (FFT) y_k^1 of the time series is obtained and multiplied by $f^{-k/2}$ ($k > 0$), yielding y_k^2 ;

3) y_k^2 is symmetrized so as to obtain a real function. The pertinent inverse FFT is obtained after rounding off and truncation. The ensuing time series x_i has the desired power-spectrum properties and, by construction, is representative of non-Gaussian noises. In this work we consider noises in the range $0 \leq k \leq 2.5$, with $\Delta k = 0.25$.

Fractional Brownian motion (fBm) and fractional Gaussian noise (fGn): fBm is the only family of processes which is Gaussian, self-similar, and endowed with stationary increments (see Ref. [55] and references therein). The normalized family of these Gaussian processes, $\{B^{\mathcal{H}}(t), t > 0\}$, has the following properties: *i)* $B^{\mathcal{H}}(0) = 0$ almost surely, *ii)* $\mathbb{E}[B^{\mathcal{H}}(t)] = 0$ (zero mean), and *iii)* covariance given by

$$\mathbb{E}[B^{\mathcal{H}}(t)B^{\mathcal{H}}(s)] = (t^{2\mathcal{H}} + s^{2\mathcal{H}} - |t - s|^{2\mathcal{H}}) / 2, \quad (2)$$

for $s, t \in \mathbb{R}$. Here $\mathbb{E}[\cdot]$ refers to the mean. The power exponent $0 < \mathcal{H} < 1$ is commonly known as the Hurst parameter or Hurst exponent. These processes exhibit *memory* for any Hurst parameter except for $\mathcal{H} = 1/2$, as one realizes from Eq. (2). The $\mathcal{H} = 1/2$ case corresponds to classical Brownian motion and successive motion-increments are as likely to have the same sign as the opposite (there is no correlation among them). Thus, Hurst's parameter defines two distinct regions in the interval $(0, 1)$. When $\mathcal{H} > 1/2$, consecutive increments tend to have the same sign so that these processes are *persistent*. On the contrary, for $\mathcal{H} < 1/2$, consecutive increments are more likely to have opposite signs *anti-persistent*.

Let us introduce the quantity $\{W^{\mathcal{H}}(t), t > 0\}$ (fBm-“increments”)

$$W^{\mathcal{H}}(t) = B^{\mathcal{H}}(t+1) - B^{\mathcal{H}}(t), \quad (3)$$

so as to express our Gaussian noise in the fashion

$$\begin{aligned} \rho(k) &= \mathbb{E}[W^{\mathcal{H}}(t)W^{\mathcal{H}}(t+k)] \\ &= \frac{1}{2} [(k+1)^{2\mathcal{H}} - 2k^{2\mathcal{H}} + |k-1|^{2\mathcal{H}}], \quad k > 0. \end{aligned} \quad (4)$$

Note that for $\mathcal{H} = 1/2$ all correlations at nonzero lags vanish and $\{W^{1/2}(t), t > 0\}$ represents *white Gaussian noise*.

The fBm and fGn are continuous but non-differentiable processes (in the classical sense). As non-stationary processes, they do not possess a spectrum defined in the usual sense; however, it is possible to define a *generalized power spectrum* of the form:

$$\Phi \propto |f|^{-\theta}, \quad (5)$$

with $\theta \equiv \alpha = 2\mathcal{H} + 1$, and $1 < \alpha < 3$ for fBm and; $\theta \equiv \beta = 2\mathcal{H} - 1$, and $-1 < \beta < 1$, for fGn.

We use the Matlab function “wfbm” that returns a fractional Brownian motion signal with a Hurst parameter \mathcal{H} ($0 < \mathcal{H} < 1$) and length N for the generation the fBm and fGn time series. The algorithm was proposed by Abry and Sellan [56, 57]. In this work we consider noises in the range $0.1 \leq \mathcal{H} \leq 0.9$.

2.2 Horizontal Visibility Graph

The horizontal visibility graph (HVG) [23] is a geometrical simplification of the visibility graph (VG) [19] that maintains the inherent characteristics of the transformed time series and incorporates in a natural way its time causality.

By construction, the HVG transforms a time series into a graph, in which each node corresponds to a point in the time series, will be connected considering the following criterion:

Let $\{x_i, i = 1, \dots, N\}$, be a time series of N data. Two nodes i and j in the graph are connected if it is possible to trace a horizontal line in the time series linking x_i and x_j not intersecting intermediate data height, fulfilling: $x_i, x_j > x_n$ for all $i < n < j$.

Note that the HVG preserves the time causality of the original series where each node sees at least its nearest neighbors. Another important feature of the HVG is the invariance under affine transformations, as its visibility is not modified under rescaling of horizontal and vertical axes, as well as under horizontal and vertical translations. Some other interesting properties are discussed in [23]. An example of a time series and its associated node degree distribution based on HVG is given in Figure 1.

2.2.1 The λ rule

Lacassa and Toral [22] propose that chaotic and stochastic time series map into a graph with an exponential node degree distribution $P(\kappa) \sim \exp(-\lambda\kappa)$. The λ parameter is computed by adjusting, using the least square method, a straight line being λ its slope. The linear scaling region considered by Lacassa and Toral is $3 \leq \kappa \leq 20$ or $3 \leq \kappa \leq \kappa_{\max}$ (if $\kappa_{\max} \leq 20$) for stochastic processes; and $3 \leq \kappa \leq 25$ or $3 \leq \kappa \leq \kappa_{\max}$ (if $\kappa_{\max} \leq 25$) for chaotic ones. The parameter λ characterizes chaotic processes when $\lambda < \ln(3/2)$, uncorrelated noises for $\lambda = \ln(3/2)$ and correlated noises when $\lambda > \ln(3/2)$.

In the same fashion, we have computed the node degree distribution $P(\kappa)$ of the HVG for all the systems described in Section 2.1 for series of 10^5 data length. Symmetric confidence intervals at the 95% confidence level were obtained assuming the Gaussian model, a linear structure for the regression and independent zero-mean errors. The results are shown in Figures 2 and 3 for all studied chaotic and stochastic dynamics, respectively. It is possible to see from these figures that several chaotic and stochastic systems follow the above mentioned rule; however, we have found others that do not. Examples of chaotic maps for which λ is larger than λ_c (see Fig. 2 open circles) correspond to: (5) cubic map, (9) Pinchers map, (10) Spence map, (11) sine-circle map, (14) delay logistic map, (15) Tinkerbell map (X), (16) Burger's map (Y), (17) Holmes cubic map, (19) Ikeda map (Y) (21) discrete predator-prey map (Y), (23) Hénon area-preserving quadratic map, (26) chaotic web map, (27) Lorenz three-dimensional chaotic map, (see also Table S1 in supplementary material). Stochastic processes for which λ is smaller than λ_c (see Fig. 3 open circles) correspond to fGn with $-0.8 \leq \beta \leq -0.4$.

Some important issues to be discussed are:

Scaling zone: Several systems present a well defined linear scaling region allowing a good linear fitting to obtain λ . Examples are the Logistic map, Holmes cubic map (X), a k -noise with $k = 0$ and a fBm with $\beta = 0$ presented in Figure 4. However, we must point out that the fact of having a well scaling region does not guarantee the satisfaction of the λ rule. See for instance the Holmes cubic map (X) that present a clear linear scaling region, however $\lambda = 0.443$ $[0.418; 0.469] > \lambda_c$, contradicting the hypothesis.

Another important point is the selection of the scaling zone, as the inclusion or exclusion of a few points in the extremes of the PDF may drastically change the λ value. Figure 5 shows the effect of selecting different scaling zones for a stochastic process with $f^{-1.75}$ PS. If the scaling zone is defined in the node degree interval $3 \leq \kappa \leq 14$, $\lambda = 0.803$ $[0.666; 0.941]$, however, if the scaling

zone is redefined for the interval $7 \leq \kappa \leq 12$, $\lambda = 0.966$ [0.797; 1.136], which represent a variation of 16.3 %.

Heavytailedness: The definition of an unique linear scaling zone is a difficult task for systems with a heavy tailed PDF. Note that, when defining a scaling zone, important information contained in the tails may be lost. Examples of systems with heavy tailed PDFs are the Cusp and the Schuster maps (see Fig. 6).

Nonexponential behavior: Some systems present PDFs with no linear scaling zone, in consequence, the hypothesis of an exponential behavior cannot be confirmed. See for instance, the Tinkerbell map (Y) and the Burger's map (X) in Fig. 6.

In Table S1 readers find the λ values with the corresponding confidence intervals, the coefficient of determination R^2 , as well all the corresponding plots for all the dynamical systems analyzed in this work (see Figures S4-S9).

2.2.2 Skewness and kurtosis

Given a one-dimensional probability distribution $f(x)$ with $x \in \Delta \subset \mathbb{R}$, the usual spread measure is the variance $V[f] = \int_{\Delta} (x - \langle x \rangle)^2 dx$. The variance measures the (quadratic) variability around the mean. This property makes the variance (or its square root, the standard deviation) particularly useful for smooth unimodal distributions. Other interesting quantifiers based on higher moments order are the skewness (a third order moment measure) and the kurtosis (which depends on the fourth order moment). The skewness measures the asymmetry, while the kurtosis describes the relative “peakedness” of the density with respect to the Gaussian law. Kurtosis is a sign of “flattening” or “peakedness” of a distribution.

The usual skewness and kurtosis are of limited use and interpretability when dealing with asymmetric distributions, as is the case of the node degree distribution HVG-PDF $P(\kappa)$, which is always non-negative.

Among the many alternatives available in the literature, for skewness and kurtosis evaluation, Brys *et al.* [58] employ with success the information provided by the quantiles. In particular, we will see that an alternative measure of kurtosis is able to describe the different heavytailedness of the observed node degree distribution HVG-PDF.

Consider $\mathbf{x} = (x_1, \dots, x_n)$ a sample of n real values. The sample quantile of order $0 < \eta < 1$ is $q(\eta) = \min\{x \in \mathbb{R} : \hat{F}(x) \geq \eta\}$ and $\hat{F}(x) = n^{-1} \#\{x_i : x_i \leq x\}$ is the sample cumulative distribution function also known as empirical function. Quantile-based measures of skewness and kurtosis can be defined as

$$\hat{\gamma}_1(\xi) = \frac{q(1 - \xi) + q(\xi) - 2q(1/2)}{q(1 - \xi) - q(\xi)}, \quad (6)$$

and

$$\hat{\gamma}_2(\xi, \varrho) = \frac{q(1 - \varrho) - q(\varrho)}{q(1 - \xi) - q(\xi)} \quad (7)$$

respectively, where $0 < \xi < \varrho < 1$ are arbitrary.

The values for two reference distributions were computed analytically with $\xi = 1/10$ and $\varrho = 1/100$. For the standard exponential distribution with probability density function $P(\kappa) = e^{-\kappa}$, $\kappa > 0$ they are:

$$\gamma_1(1/10) = \frac{2 \log(1/2) - \log(1/10) - \log(9/10)}{\log(9/10) - \log(1/10)} \approx 0.465, \quad (8)$$

and

$$\gamma_2(1/10, 1/100) = \frac{\log(99/100) - \log(1/100)}{\log(9/10) - \log(1/10)} \approx 2.091, \quad (9)$$

and for the node degree distribution under white noise, whose probability function is $P(k) = 3^{-1}(2/3)^{k-2}$, $k \geq 2$ [19, 22], they are $\gamma_1(1/10) = 3/5 = 0.6$ and $\gamma_2(1/10, 1/100) = 11/5 = 2.2$.

Table 1 shows the values of lambda ($\hat{\lambda}$), skewness ($\hat{\gamma}_1$) and kurtosis ($\hat{\gamma}_2$) for several noises and chaotic maps.

It is worth noticing that several chaotic maps present high kurtosis values indicating a heavy tailed PDF, showing the importance of using a quantifier that considers the entire available data. Fig. 7 displays examples of HVG-PDF of several chaotic and stochastic systems. Note that, for some systems, the HVG-PDFs do not present an exponential behavior. Readers can find the results for all the systems considered in Table S1 of the supplementary material.

2.3 The Shannon-Fisher information plane

To avoid the subjectivity of choosing the scaling zone in which the parameter λ is computed and, consequently, the sensitivity of this methodology, we propose a tool in which no information is lost, as the entire PDF is used and the relation between global and local features of the systems is captured. The Shannon-Fisher information plane ($\mathcal{S} \times \mathcal{F}$) firstly introduced by Vignat and Bercher [59] is a planar representation in which the horizontal and vertical axes are functionals of the pertinent probability distribution, namely, the Shannon Entropy \mathcal{S} and the Fisher Information measure \mathcal{F} , respectively. This tool is a convenient way to represent in the same information plane global and local aspects of the PDFs associated to the studied system. In this work the PDFs are obtained through the horizontal visibility graph methodology [19].

Given a continuous probability distribution function (PDF) $f(x)$ with $x \in \Delta \subset \mathbb{R}$ and $\int_{\Delta} f(x) dx = 1$, its *Shannon Entropy* [60] is

$$\mathcal{S}[f] = - \int_{\Delta} f \ln f dx, \quad (10)$$

a measure of “global” character that it is not too sensitive to strong changes in the distribution taking place on small regions of the support Δ .

Such is not the case with *Fisher’s Information Measure* (FIM) \mathcal{F} [61–63], which constitutes a measure of the gradient content of the distribution $f(x)$, thus being quite sensitive even to tiny localized perturbations. It reads

$$\mathcal{F}[f] = \int_{\Delta} \frac{1}{f(x)} \left[\frac{df(x)}{dx} \right]^2 dx = 4 \int_{\Delta} \left[\frac{d\psi(x)}{dx} \right]^2 dx. \quad (11)$$

FIM can be variously interpreted as a measure of the ability to estimate a parameter, as the amount of information that can be extracted from a set of measurements, and also as a measure of the state of disorder of a system or phenomenon [63]. In the previous definition of FIM (Eq. (11)) the division by $f(x)$ is not convenient if $f(x)$ becomes too small to be adequately computed. Such issue is avoided using probability amplitudes \sqrt{f} [62, 63]. The gradient operator significantly influences the contribution of minute local f -variations to FIM’s value. Accordingly, this quantifier is called a “local” one [63].

Let now $P = \{p_i; i = 1, \dots, N\}$ be a discrete probability distribution, with N the number of possible states of the system under study. The concomitant problem of information-loss due to discretization has been thoroughly studied (see, for instance, [64–66], and references therein) and, in particular, it entails the loss of FIM’s shift-invariance, which is of no importance for our present purposes [17, 18]. In the discrete case, we define a “normalized” Shannon entropy as

$$H[P] = \mathcal{S}[P]/S_{\max} = \left\{ - \sum_{i=1}^N p_i \ln p_i \right\} / S_{\max}, \quad (12)$$

where the denominator $S_{\max} = \mathcal{S}[P_e] = \ln N$ is the Shannon entropy attained by a uniform probability distribution $P_e = \{p_i = 1/N, \forall i = 1, \dots, N\}$. For the FIM we take the expression in terms of real

probability amplitudes as starting point, then a discrete normalized FIM convenient for our present purposes, is given by

$$\mathcal{F}[P] = F_0 \sum_{i=1}^{N-1} [(p_{i+1})^{1/2} - (p_i)^{1/2}]^2. \quad (13)$$

It has been extensively discussed that this discretization is the best behaved in a discrete environment [67]. Here the normalization constant F_0 reads

$$F_0 = \begin{cases} 1 & \text{if } p_{i^*} = 1 \text{ for } i^* = 1 \text{ or } i^* = N \text{ and } p_i = 0 \ \forall i \neq i^* \\ 1/2 & \text{otherwise.} \end{cases} \quad (14)$$

3 Results and discussion

In order to study the stability of the forthcoming results, we first analyze the dependency of the Information Theory quantifiers with the size of the time series. For this experiment we consider time series with different length sizes, varying from 30,000 to 500,000 values. As it can be seen in Figure 8, the Fisher Information and the Shannon entropy rapidly converge to stable values. For example, for the cases depicted in Figure 8, the order of magnitude of the percentage variations of the mean value for times series with $N = 100,000$ and $N = 500,000$ are between 10^{-4} and 10^{-5} . For that reason all experiments consider time series with 100,000 values. The Fisher Information and the Shannon Entropy are computed for all systems presented in Section 2.1. Results are depicted in Figures 9, 10 and 11. The Shannon entropy values are normalized with its maximum value for $N = 100,000$, that corresponds to the entropy of the gaussian white noise (fBm for $\beta = 0$, $\mathcal{S}_{[P_{wgn}]}$). One interesting observation resulting from this experiment is the fact that the Fisher Information (\mathcal{F}) decreases with the strength of correlation in noises. The degree distribution corresponding to noises with lower correlation presents high peaks as well as long tails, almost flat for the white noise. As correlations get stronger, the peaks decrease and tails get shorter. For the uncorrelated situation (white noise), the strong contribution of the long and flat tail, even having the highest peak, makes the shape of the distribution more uniform. This effect can be seen in Figure 7 as well as in Table S1 for noises with f^{-k} power spectrum.

The Fisher Information is sensitive to small fluctuations. From Eq. (13), it is possible to see that bigger differences in consecutive p_i values of the distribution P , result in higher values of \mathcal{F} . In this case, the higher peaks in the degree distributions, that correspond to lower correlation values, represent the main contribution to \mathcal{F} . The extra terms present in the long tail, even contributing with small values, still increase the value of \mathcal{F} . For that reason, the lowest value of \mathcal{F} corresponds to the noise with strongest correlation structure (fBm with $\alpha = 2.8$), see Figure 11.

That is not the case of the Shannon entropy \mathcal{S} , which is not sensitive to small fluctuations. The Shannon Entropy presents its highest value for noises with the smallest correlation (white noises, gaussian and non-gaussian) and, as correlation structures get stronger, \mathcal{S} decreases.

The planar localization in the Shannon-Fisher information plane $\mathcal{S} \times \mathcal{F}$, gives interesting information about the relation between the systems. Noises appear to be organized as a frontier, from which all chaotic maps concentrate. As it was previously shown, the frontier is stable regarding the size of the times series length.

Note that some chaotic maps are located nearby the noise “frontier” in the $\mathcal{S} \times \mathcal{F}$ plane (see Figure 11). These maps are: the linear congruential generator (4), the dissipative standard map (18), and the Sinai map (20). They present high $\mathcal{S}/\mathcal{S}_{wgn}$ values, $0.995 \leq \mathcal{S}/\mathcal{S}_{wgn} \leq 1$ and low \mathcal{F} values, $0.17 \leq \mathcal{F} \leq 0.19$. This planar localization can be understood, as these maps present a stochastic like dynamical behavior when represented in a two dimensional plane. However, when represented in higher dimensional planes, planar structures appear denoting their chaotic behavior.

The use of the $\mathcal{S} \times \mathcal{F}$ plane can shed light on the underlying system’s structure. For example, the Schuster maps display a linear behavior in the $\mathcal{S} \times \mathcal{F}$ plane when varying the z parameter. Wider

laminar regions ($z = 2.00$) generate a greater number of nodes with lower degree values. At the same time nodes located in the extremes of a laminar region possess higher degree. As the parameter z decreases, the laminar structures get thinner, reducing the number of nodes with higher degree value. This fact positioned the Schuster systems far from the frontier as z increases as can be seen in Figure 9. In Table S1 in the supplementary material can be found a detailed description of the results.

4 Conclusions

Time series (TS), temporal sequences of measurements or observations, are the basic elements for investigating natural phenomena. From TS one should judiciously extract information of dynamical systems. TS arising from chaotic systems share with those generated by stochastic processes several properties that make them very similar. Examples of these properties are: a wide-band power spectrum (PS), a delta-like autocorrelation function, and an irregular behavior of the measured signals. As irregular and apparently unpredictable behavior is often observed in natural TS, the question that immediately emerges is whether the system is chaotic (low-dimensional deterministic) or stochastic. If one is able to show that the system is dominated by low-dimensional deterministic chaos, then only few (nonlinear and collective) modes are required to describe the pertinent dynamics [1]. If not, then, the complex behavior could be modeled by a system dominated by a very large number of excited modes which are in general better described by stochastic or statistical approaches.

In the present contribution, we propose the use of the Horizontal Visibility Graph in combination with the Shannon entropy and the Fisher information measure as a methodology to study dynamical systems. Several chaotic and correlated noises were considered for an exhaustive analysis. The arrangement of the results in the $\mathcal{S} \times \mathcal{F}$ plane, show that this novel tool is able to capture features that reveal the nature governing the system.

Summarizing, we have presented extensive series of numerical evidence and have contrasted the characterizations of deterministic chaotic and noisy-stochastic dynamics, as represented by time series of finite length. Surprisingly enough, one just has to look at the different planar locations of the two dynamical regimes. The planar location is able to tell us whether we deal with chaotic or stochastic time series.

Acknowledgments

M.G.R. acknowledges support from CNPq and FAPEMIG, Brazil. L.C.C. acknowledges support from CNPq, Brazil. O.A.R. acknowledges support from Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina, and FAPEAL, Brazil. B. Amin Gonçalves acknowledges support from CAPES, Brazil. A.C.F. acknowledges support from CNPq and FAPEAL, Brazil.

References

1. Osborne AR, Provenzale A (1989) Finite correlation dimension for stochastic systems with power-law spectra. *Physica D* 35: 357 – 381.
2. Sugihara G, May RM (1983) Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature* 344: 734 – 741.
3. Kaplan DT, Glass, L (1992) Direct test for determinism in a time series. *Phys. Rev. Lett.* 68: 427 – 430.
4. Kaplan DT, Glass L (1993) Coarse-grained embeddings of time series: random walks, Gaussian random processes, and deterministic chaos. *Physica D* 64: 431 – 454.

5. Kantz H, Olbrich E (2000) Coarse grained dynamical entropies: Investigation of high-entropic dynamical systems. *Physica A* 280: 34 – 48.
6. Cencini M, Falcioni M, Olbrich E, Kantz H, Vulpiani A (2000) Chaos or noise: Difficulties of a distinction. *Phys. Rev. E* 62: 427 – 437.
7. Grasberger P, Procaccia I (1983) Characterization of Strange Attractors. *Phys. Rev. Lett.* 50: 346 – 349.
8. Abarbanel HDI (1996) *Analysis of Observed Chaotic Data*. New York, USA: Springer-Verlag.
9. Kolmogorov, AN (1959) A new metric invariant for transitive dynamical systems and automorphisms in Lebesgue spaces. *Dokl. Akad. Nauk., USSR*, 119: 861 – 864.
10. Sinai YG (1959) On the concept of entropy for a dynamical system. *Dokl. Akad. Nauk., USSR*, 124: 768 – 771.
11. Rosso OA, Larrondo HA, Martín MT, Plastino A, Fuentes MA (2007) Distinguishing noise from chaos. *Phys. Rev. Lett.* 99: 154102.
12. Rosso OA, Carpi LC, Saco PM, Gómez Ravetti M, Plastino A, Larrondo HA (2012) Causality and the entropy-complexity plane: robustness and missing ordinal patterns. *Physica A* 391: 42 – 55.
13. Rosso OA, Carpi LC, Saco PM, Gómez Ravetti M, Larrondo HA, Plastino A (2012) The Amigó paradigm of forbidden/missing patterns: a detailed analysis. *Eur. Phys. J. B* 85: 419 – 431.
14. Rosso OA, Olivares F, Zunino L, De Micco L, Aquino ALL, Plastino A, Larrondo HA (2013) Characterization of chaotic maps using the permutation Bandt-Pompe probability distribution. *Eur. Phys. J. B* 86: 116 – 129.
15. Bandt C, Pompe B (2002) Permutation entropy: a natural complexity measure for time series. *Phys. Rev. Lett.* 88: 174102 (2002).
16. Zanin M, Zunino L, Rosso OA, Papo D (2012) Permutation Entropy and Its Main Biomedical and Econophysics Applications: A Review. *Entropy* 14: 1553 – 1577.
17. Olivares F, Plastino A, Rosso OA (2012) Ambiguities in BandtPompe’s methodology for local entropic quantifiers. *Physica A* 391: 2518 – 2526.
18. Olivares F, Plastino A, Rosso OA (2012) Contrasting chaos with noise via local versus global information quantifiers. *Phys. Lett A* 376: 1577 – 1583.
19. Lacasa L, Luque B, Ballesteros F, Luque J, Nuno JC (2008) From time series to complex networks: The visibility graph. *Proc. Natl. Acad. Sci. USA* 105: 4972 – 4975.
20. B. Luque, L. Lacasa, F. J. Ballesteros, and A. Robledo. Feigenbaum graphs: A complex network perspective of Chaos. *PLoS ONE* **6**, (2011) e22411.
21. Luque B, Lacasa L, Ballesteros FJ, Robledo A (2012) Analytical properties of horizontal visibility graphs in the Feigenbaum scenario. *Chaos* 22: 013109.
22. Lacasa L, Toral R (2010) Description of stochastic and chaotic series using visibility graphs. *Phys. Rev. E.* 82: 036120.
23. Luque B, Lacasa L, Ballesteros F, Luque J (2009) Horizontal visibility graphs: exact results for random time series, *Phys. Rev. E* 80: 046103.

24. Sprott JC (2012) Chaos and Time series analysis. Complex Systems 21: 193 – 200.
25. Schuster HG (1988) Deterministic Chaos, 2nd ed. Weinheim: VHC,
26. May R (1976) Simple mathematical models with very complicated dynamics. Nature 261: 45 – 67.
27. Strogatz SH (1994) Nonlinear dynamics and chaos with applications to physics, biology, chemistry, and engineering. Reading, MA: Addison-Wesley-Longman.
28. Devaney RL (1989) An introduction to chaotic dynamical systems, 2nd Edition. Redwood City, CA: Addison-Wesley.
29. Knuth, DE (1997) Sorting and searching , Vol. 3 of The art of computer programming, (3rd Edition), Reading, MA: Addison-Wesley-Longman.
30. Zeng W, Ding M, Li J (1985) Symbolic description of periodic windows in the antisymmetric cubic map. Chinese Physics Letter 2: 293 – 296.
31. Ricker W (1954) Stock and recruitment (1954) Journal of the Fisheries Research Board of Canada 11, 559 – 663..
32. van Wyk MA, Steeb W (1997), Chaos in electronics Dordrecht, Belgium: Kluwer.
33. Beck C, Schlögl F (1995) Thermodynamics of chaotics systems. New York, USA: Cambridge University Press.
34. Potapov A, Ali MK (2000) Robust chaos in neural networks. Phys. Lett. A 277: 310 – 322.
35. Shaw R (1981) Strange attractors, chaotic behaviour, and information flow; Zeitschrift für Naturforschung A 36: 80 – 112.
36. Arnold VI (1965) Small denominators, I: mappings of the circumference into itself. American Mathematical Society Translation Series 46:, 213 – 284.
37. Hénon M (1976) *A two-dimensional mapping with a strange attractor*, Communication on Mathematical Physics 50: 69 – 77.
38. Lozi, R (1978) Un attracteur étrange? Du type attracteur de Hénon. Journal de Physique 39, 9 – 10.
39. Aronson DG, Chory MA, Hall GR, McGehee RP (1982) Bifurcations from an invariant circle for two-parameter families of maps of the plane: a computer-assisted study, Communications in Mathematical Physics 83: 304 – 354.
40. Nusse HE, Yorke JA (1994) Dynamics: numerical explorations. New York, USA: Springer.
41. Whitehead RR, Macdonald N (1984) A chaotic mapping that displays its own homoclinic structure. Physica D 13: 401 – 407.
42. Holmes P (1979) A nonlinear oscillator with a strange attractor. Philosophical Transactions of the Royal Society of London Series A 292: 419 – 448.
43. Schmidt G, Wang BH (1985) Dissipative standard map. Physical Review A 32: 2994 – 2999.
44. Ikeda E (1979) Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system Optics Communications 30: 257 – 61.

45. Sinai Ya G (1972) Gibbs measures in ergodic theory. *Russian Mathematical Surveys* 27: 21 – 69 .
46. Beddington JR, Free CA, Lawton JH (1975) Dynamic complexity in predator-prey models framed in difference equations. *Nature* 255: 58 – 60..
47. Chirikov BV (1979) A universal instability of many-dimensional oscillator systems. *Physics Reports* 52: 273 – 379.
48. Hénon M (1969) Numerical study of quadratic area-preserving mappings. *Quarterly of Applied Mathematics* 27: 291 – 312.
49. Arnold VI, Avez A (1968) *Ergodic problems of classical mechanics*. New York, USA: Benjamin.
50. Devaney RL (1984) A piecewise linear model for the zones of instability of an area-preserving map. *Physica D* 10: 387 – 93.
51. Chernikov AA, Sagdeev RZ, Zaslavskii GM (1988) Chaos: how regular can it be? *Physics Today* 41: 27 – 35 .
52. Lorenz EN (1993) *The essence of chaos* Seattle, USA: University of Washington Press.
53. Larrondo HA (2012) *Matab program: noisefk.m* (<http://www.mathworks.com/matlabcentral/fileexchange/35381>).
54. Matsumoto M, Nishimura T (1998) Mersenne twister: a 623-dimensionally uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation* 8: 3 – 30 .
55. Zunino L, érez DG, Martín MT, Plastino A, Garavaglia M, Rosso OA (2007) Characterization of Gaussian self-similar stochastic processes using wavelet-based informational tools. *Phys. Rev. E* 75: (2007) 021115.
56. Abry P, Sellan F (1996), The wavelet-based synthesis for the fractional Brownian motion proposed by F. Sellan and Y. Meyer: Remarks and fast implementation. *Applied and Computational Harmonic Analysis* 3: 377 – 383.
57. Bardet JM, Lang G, Oppenheim G, Philippe A, Stoev S, Taqqu MS (2003) Generators of long-range dependence processes: a survey. *Theory and applications of long-range dependence*. Birkhauser, 579 – 623.
58. Brys G, Hubert M, Struyf A (2009) Robust measures of tail weight, *Computational Statistics and Data Analysis* 50: 733–759.
59. C. Vignat and J.-F. Bercher, Analysis of signals in the Fisher-Shannon information plane. *Phys. Lett. A*. **312**, (2003) 27–33.
60. Shannon CE (1948) A mathematical theory of communication, *Bell Syst. Technol. J.* 27: 379 – 423, 623 – 56 .
61. R. A. Fisher, On the mathematical foundations of theoretical statistics, *Philos. Trans. R. Soc. Lond. Ser. A*, **222** 309–368 (1922).
62. Frieden BR (1998) *Physics from Fisher Information: A Unification*. Cambridge, USA: Cambridge University Press,
63. Frieden, B. R. *Science from Fisher information: A Unification*. Cambridge, UK: Cambridge University Press (2004).

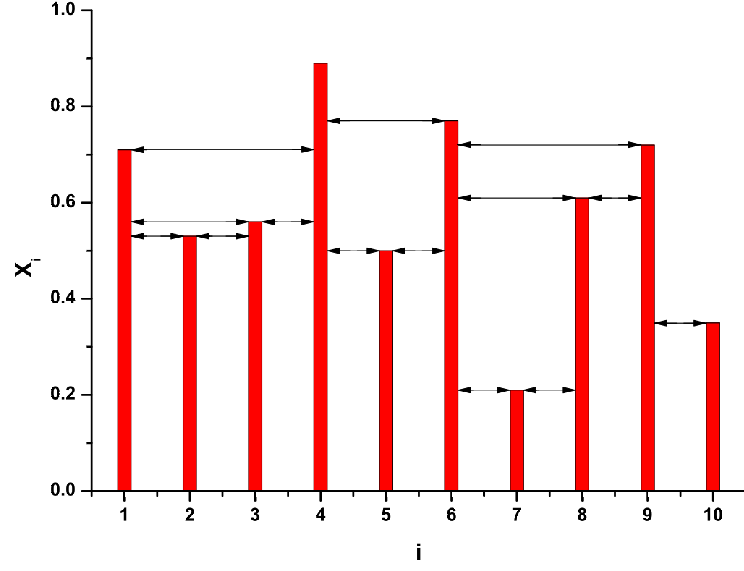
- 64. Zografos K, Ferentinos K, Papaioannou T (1986) Discrete approximations to the Csiszár, Renyi, and Fisher measures of information, *Canad. J. Stat.* 14: 355 – 366 .
- 65. Pardo L, Morales D, Ferentinos K, Zografos K (1994) Discretization problems on generalized entropies and R-divergences, *Kybernetika* 30: 445 – 460.
- 66. Madiman M, Johnson O, Kontoyiannis I (2007) Fisher Information, compound Poisson approximation, and the Poisson channel. *IEEE International Symposium on Information Theory, 2007. ISIT 2007*: 976 – 980.
- 67. Sánchez-Moreno P, Yáñez RJ, Dehesa JS (2009) Discrete Densities and Fisher Information. In *Proceedings of the 14th International Conference on Difference Equations and Applications. Difference Equations and Applications*. Istanbul, Turkey: Bahçeşehir University Press, 291 – 298.

Tables

Table 1. Dynamical systems and their statistical quantifiers skewness ($\hat{\gamma}_1$), kurtosis ($\hat{\gamma}_2$) evaluated for $\xi = 1/10$ and $\varrho = 1/100$. $\hat{\lambda}$ is the obtained parameter value (with confidence value interval) of exponential functional form proposed by Lacasa and Toral for the HVG-PDF [22]). Time series with $N = 10^5$ data are considered.

System	$\hat{\lambda}$	$\hat{\gamma}_1(\xi)$	$\hat{\gamma}_2(\xi, \varrho)$
Exponential	1	0.465	2.091
White Noise	$\log 3/2 \approx 0.405$	0.600	2.200
$k = 0$ Noise	0.404 [0.398; 0.409]	0.600	2.200
$\alpha = 2$ fBm	0.726 [0.660; 0.791]	0.200	1.600
$\beta = 0$ fGn	0.407 [0.400; 0.414]	0.600	2.200
Logistic map	0.281 [0.264; 0.298]	0.666	2.333
Cusp map	0.217 [0.175; 0.259]	0.500	5.250
$z = 1.25$ Schuster	0.387 [0.362; 0.412]	0.600	2
$z = 2$ Schuster	0.252 [0.196; 0.308]	-1	13

Figure Legends



$$\begin{aligned}
 X(t) = \{x_i; i = 1, \dots, 10\} &= \{0.71 \quad 0.53 \quad 0.56 \quad 0.89 \quad 0.50 \quad 0.77 \quad 0.21 \quad 0.60 \quad 0.72 \quad 0.35\} \\
 P(\kappa) = \{p_\kappa; \kappa = 0, \dots, 9\} &= \{0 \quad 0.1 \quad 0.3 \quad 0.4 \quad 0.1 \quad 0.1 \quad 0.0 \quad 0 \quad 0 \quad 0\}
 \end{aligned}$$

Figure 1. Horizontal Visibility Graph method applied to the time series $X(t)$. $P(\kappa)$ denotes the node degree distribution of the obtained graph (HVG-PDF).

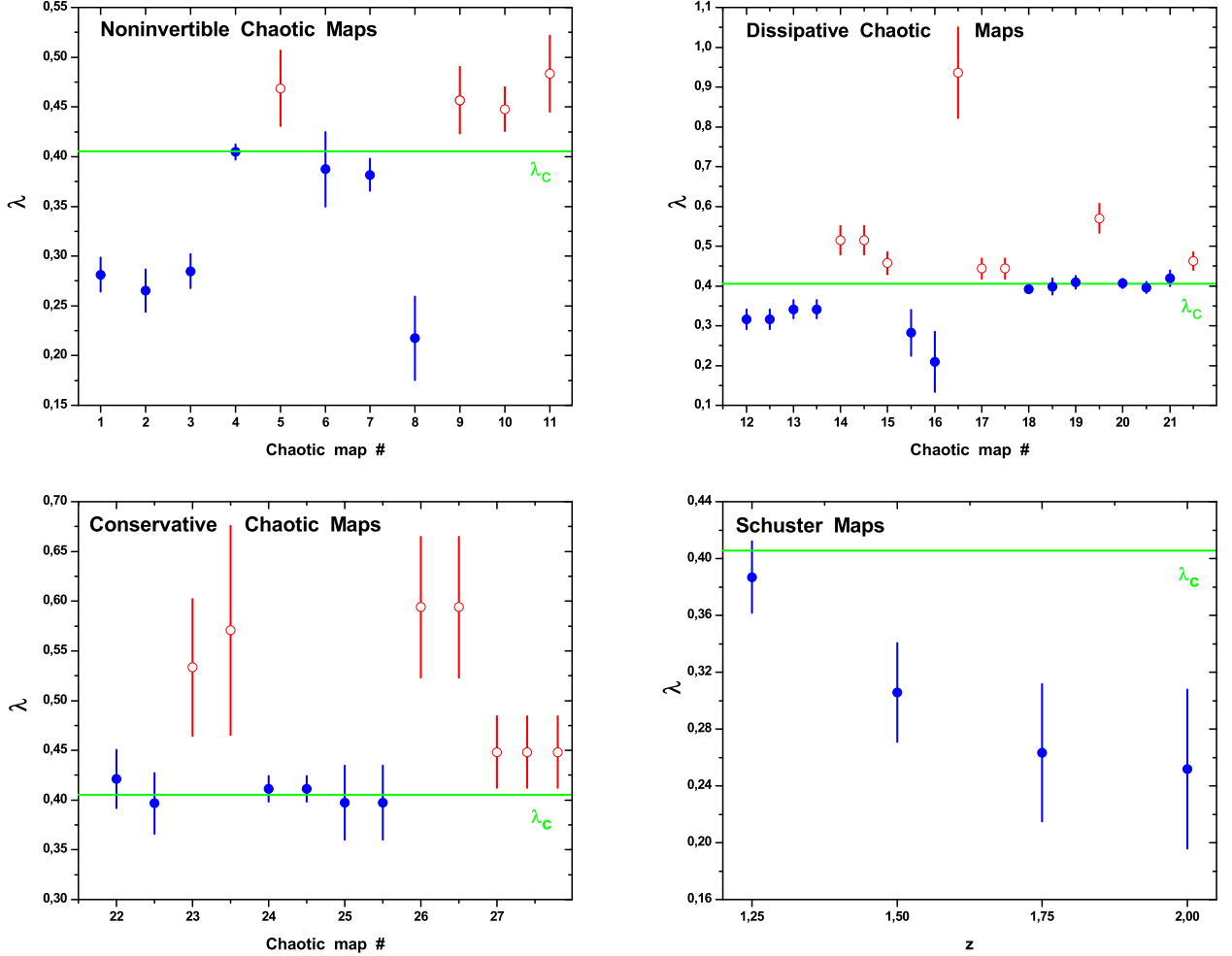


Figure 2. Confidence intervals for λ values for Noninvertible, Dissipative, Conservative and Schuster Chaotic Maps. The values were obtained following the methodology proposed by Lacasa *et al.*. Symmetric confidence intervals at the 95% confidence level were obtained for the λ parameter assuming the Gaussian model, a linear structure for the regression and independent zero-mean errors. The horizontal line represents the value of λ_c corresponding to white noise (uncorrelated stochastic dynamics). The list of names for each map is the same given in Sec. 2.1.1. Full circles (blue) are in agreement with Lacasa and Toral [22] proposal rule. Empty circles (red) not.

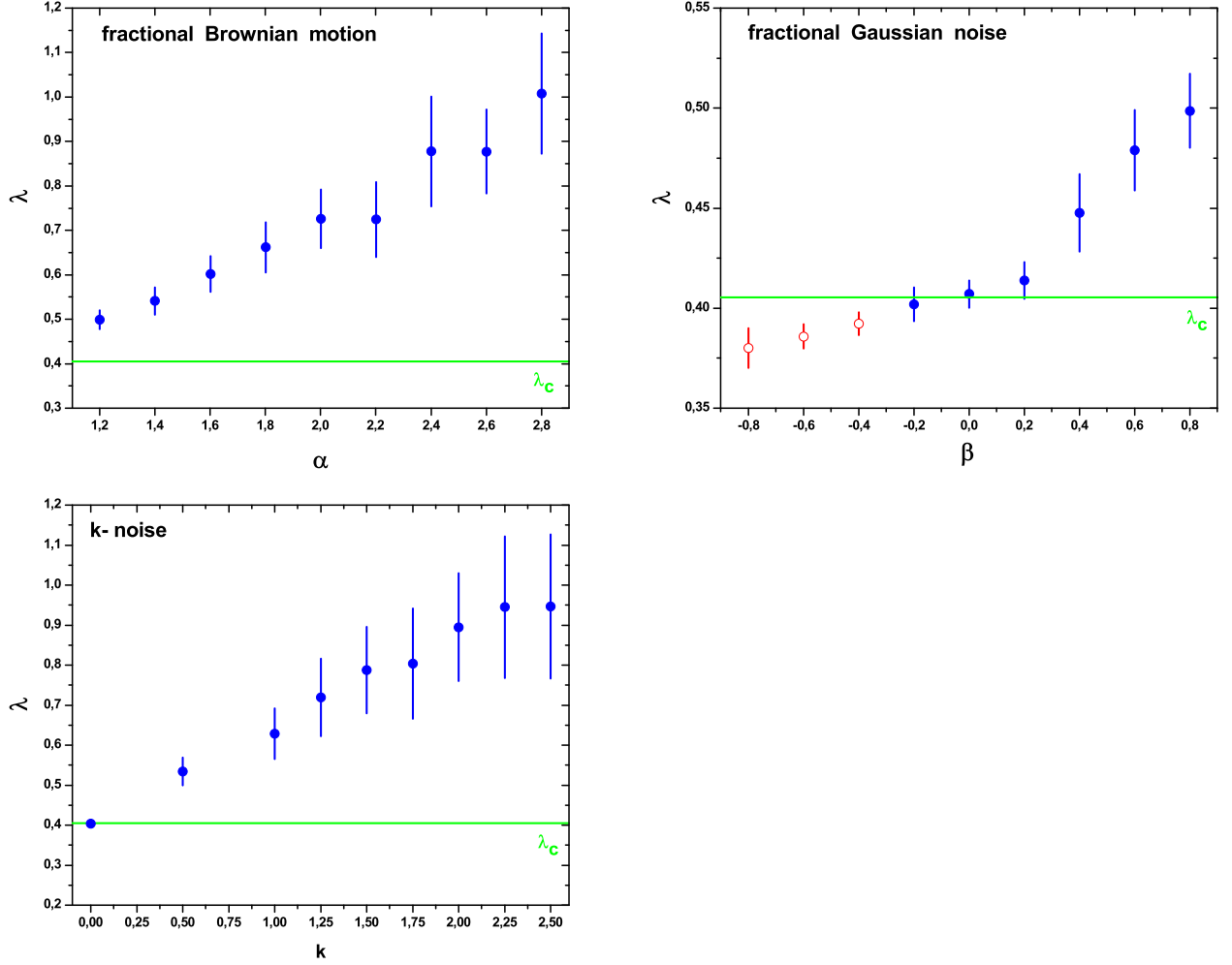


Figure 3. Parameter λ values of HVG-PDF $P(\kappa)$ for fBm, fGn and noise with f^{-k} power spectrum time series with total length of $N = 10^5$ data. The λ values were obtained following the methodology proposed by Lacasa *et al.*: from the graph $\ln[P(\kappa)]$ versus κ , the λ parameter was computed by adjusting using the least square method, a straight line being λ its slope. The linear scaling region considered in all cases is $3 \leq \kappa \leq 20$, or $3 \leq \kappa \leq \kappa_{\max}$ (if $\kappa_{\max} \leq 20$). Symmetric confidence intervals at the 95% confidence level were obtained for the λ parameter assuming the Gaussian model, a linear structure for the regression and independent zero-mean errors. The horizontal line represents the value of λ_c corresponding to white noise (uncorrelated stochastic dynamics) Full circles (blue) are in agreement with Lacassa and Toral [22] proposal rule. Empty circles (red) not.

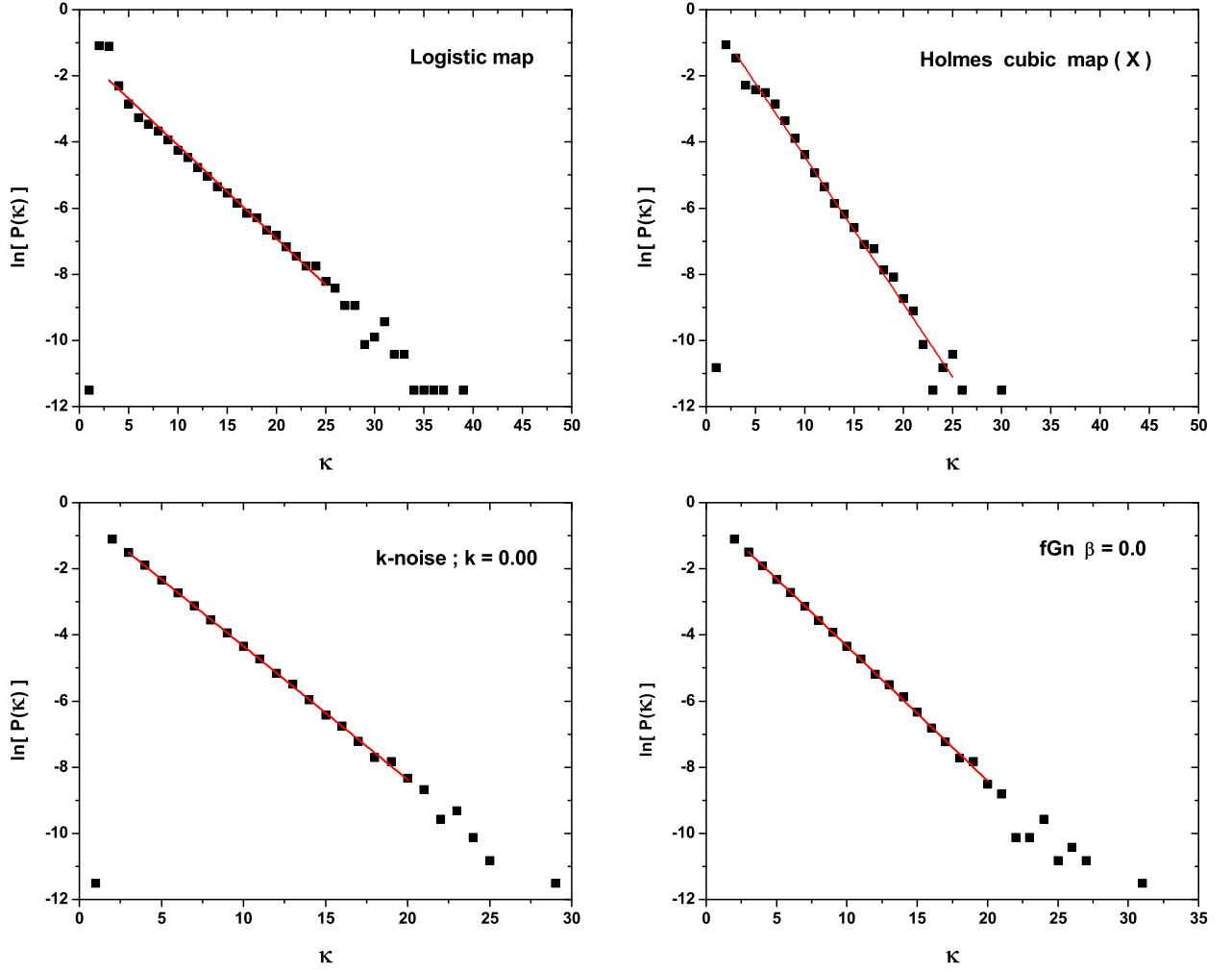


Figure 4. λ -value determination: examples of analyzed dynamical systems where a good linear scaling region was found. For the Holmes cubic map (X), however, even having a good fitting, the λ -value obtained is greater than λ_c which not satisfied the chaotic distinction suggested by Lacasa and Toral [22]. In all cases, time series with $N = 10^5$ are considered, and linear scaling regions are defined by $3 \leq \kappa \leq 25$ for chaotic and $3 \leq \kappa \leq 20$ for stochastic time series.

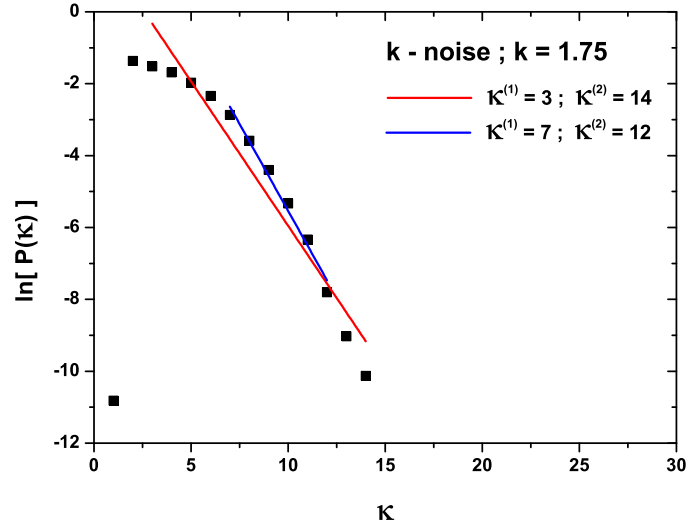


Figure 5. λ -value determination in the case of time series generated by stochastic dynamics with f^{-k} power spectrum with $k = 1,75$. Time series with $N = 10^5$ data. Two different linear scaling zones: *a*) $3 \leq \kappa \leq 14$ given $\lambda = 0.803$ $[0.666; 0.941]$; and *b*) $7 \leq \kappa \leq 12$ given $\lambda = 0.966$ $[0.797; 1.136]$. Note that the slope of the straight line change significantly.

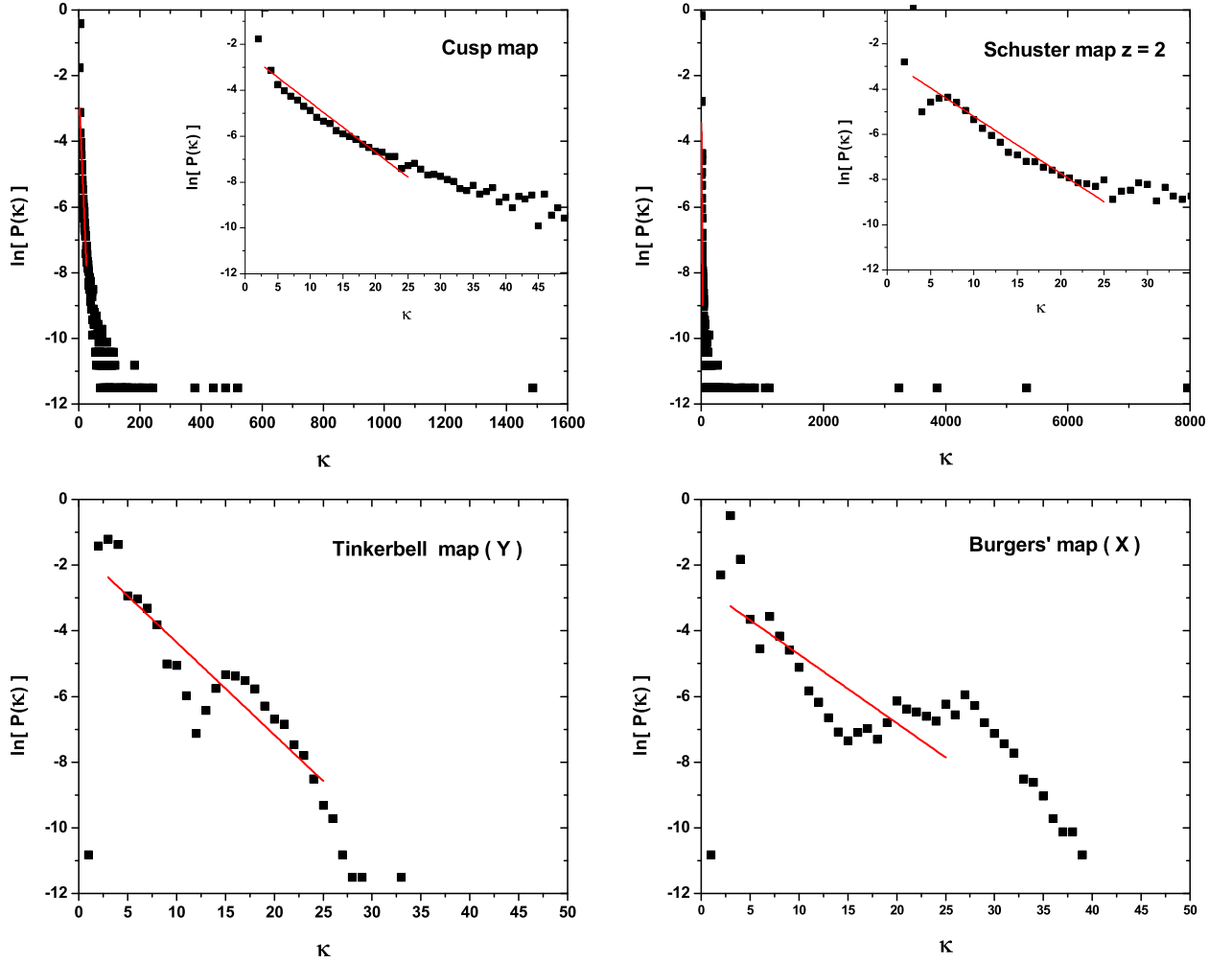


Figure 6. Cases with bad λ -value determination: *a*) Cusp map and Schuster map with $z = 2$, the associated HVG-PDF present heavy tail making difficult to define a unique linear scaling zone representative of all the data. *b*) Tinkerbell map (Y) and the Burger's map (X) for which it is impossible to define a unique linear scaling zone, and in consequence the hypothesis of an exponential behavior cannot be confirmed. Time series with $N = 10^5$ data are considered.

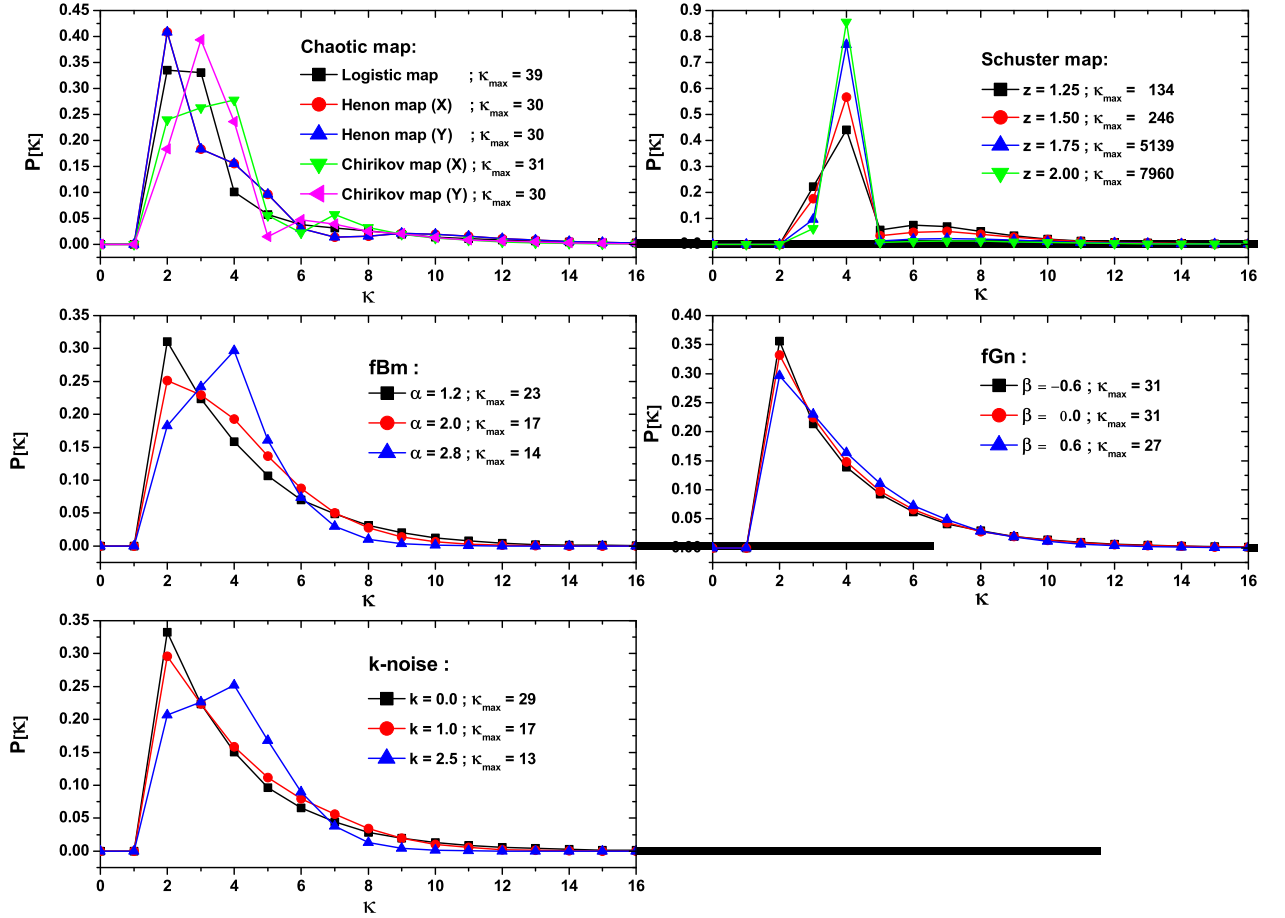


Figure 7. Examples of HVG-PDF for some chaotic and stochastic systems. Only $0 \leq \kappa \leq 16$ are displayed. Note that the corresponding cut-offs (κ_{\max}) are also shown. The length of the time series is $N = 100,000$.

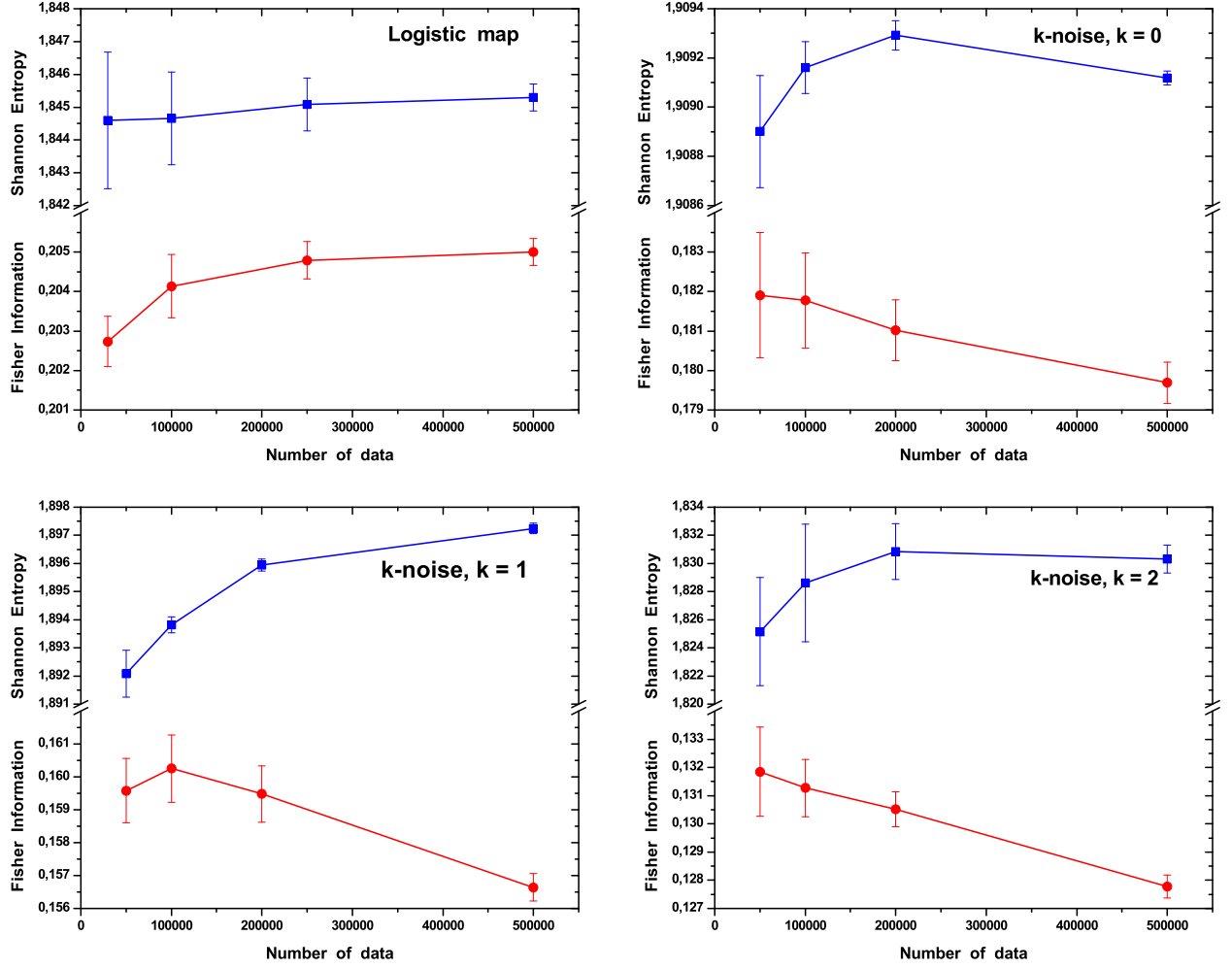


Figure 8. Study of the effect of the series length on the Information Theory quantifiers. The dynamical systems here considered are the Logistic map and noises with f^{-k} power spectrum, for $k = 0, 1$ and 2 .

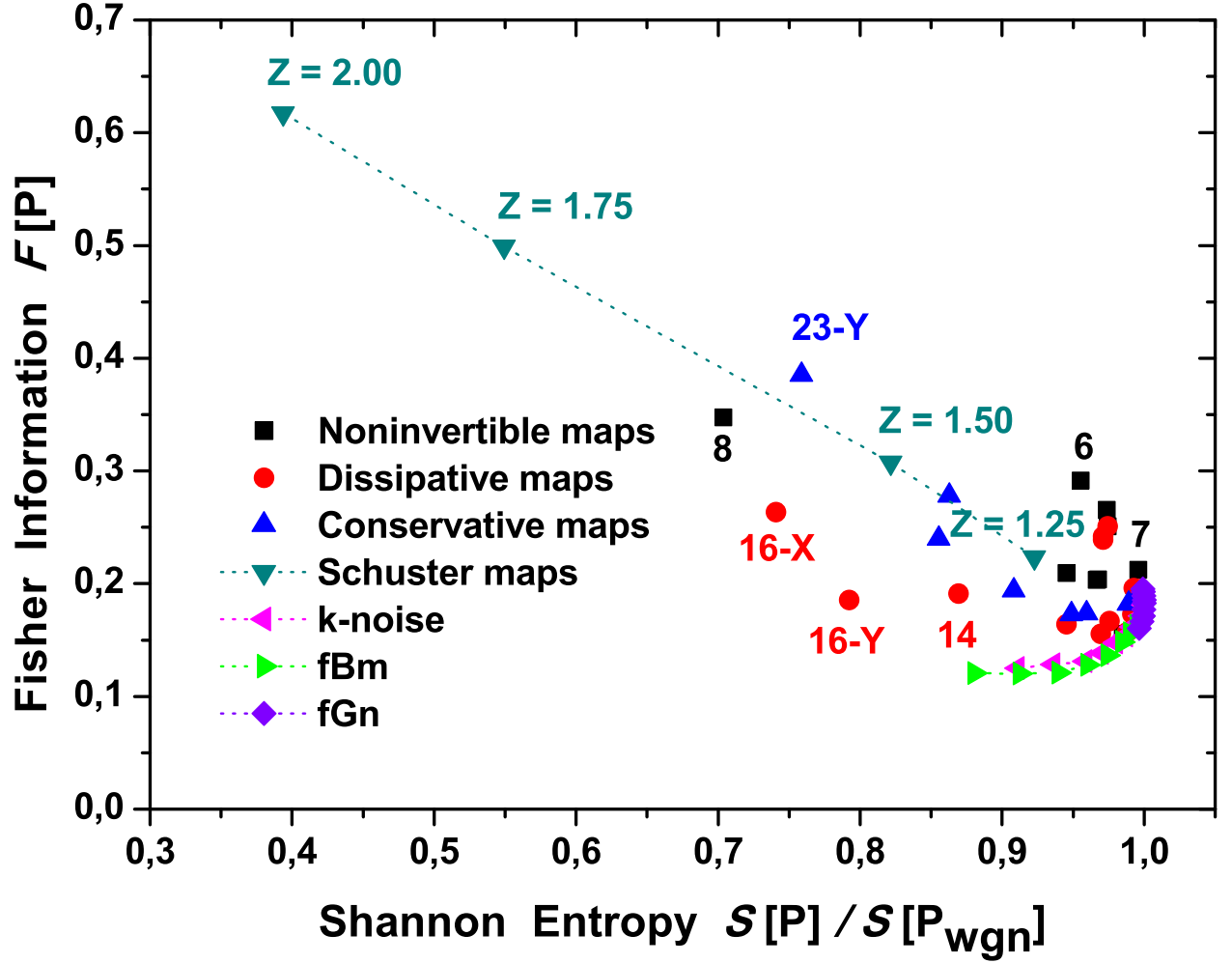


Figure 9. Representation on the Shannon-Fisher plane, $\mathcal{S} \times \mathcal{F}$, for all dynamical systems. The quantifiers were evaluated with the HVG-PDF from time series length 10^5 .

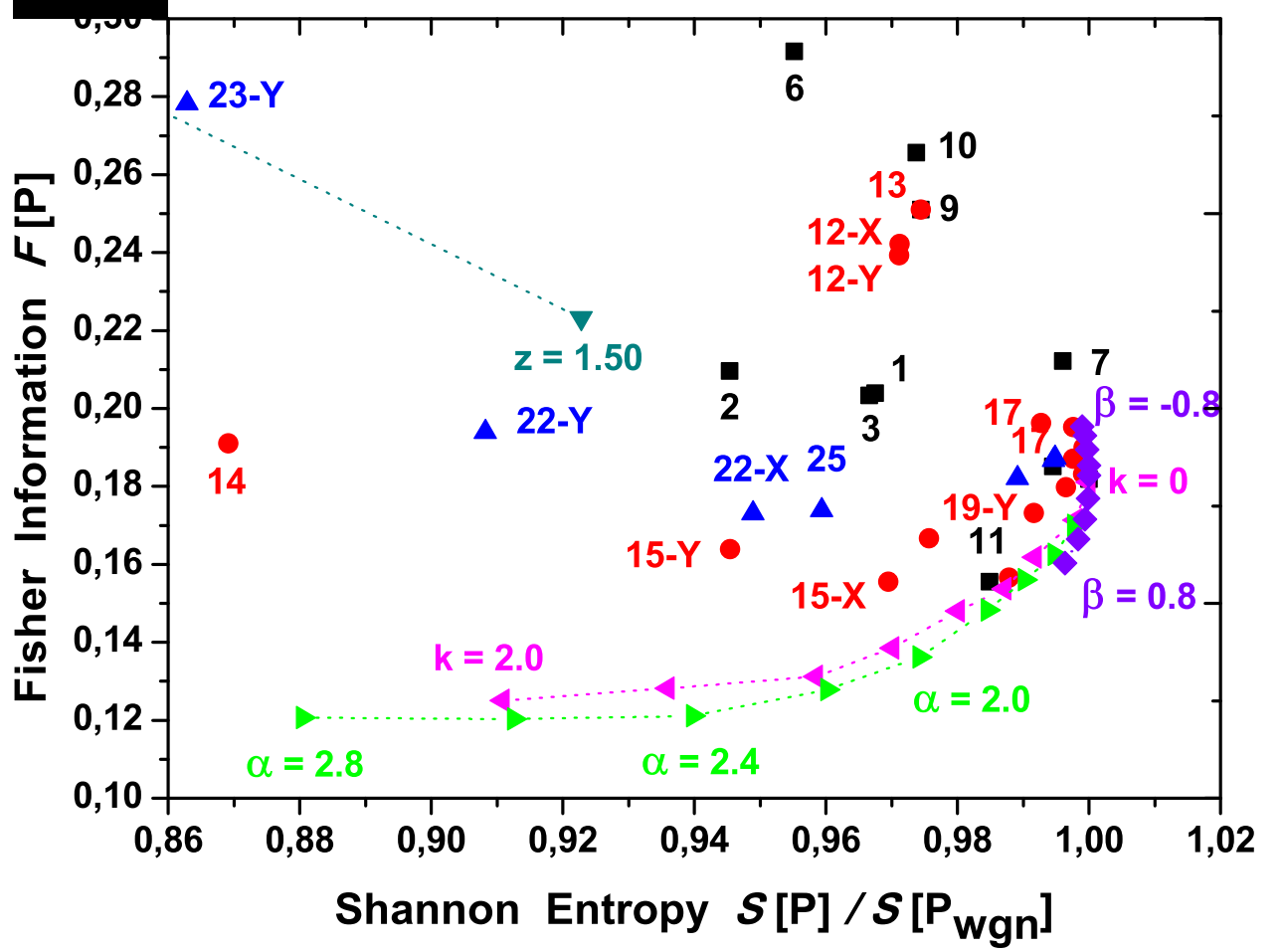


Figure 10. Shannon-Fisher plane, $S \times \mathcal{F}$ zoom, see Fig. 9

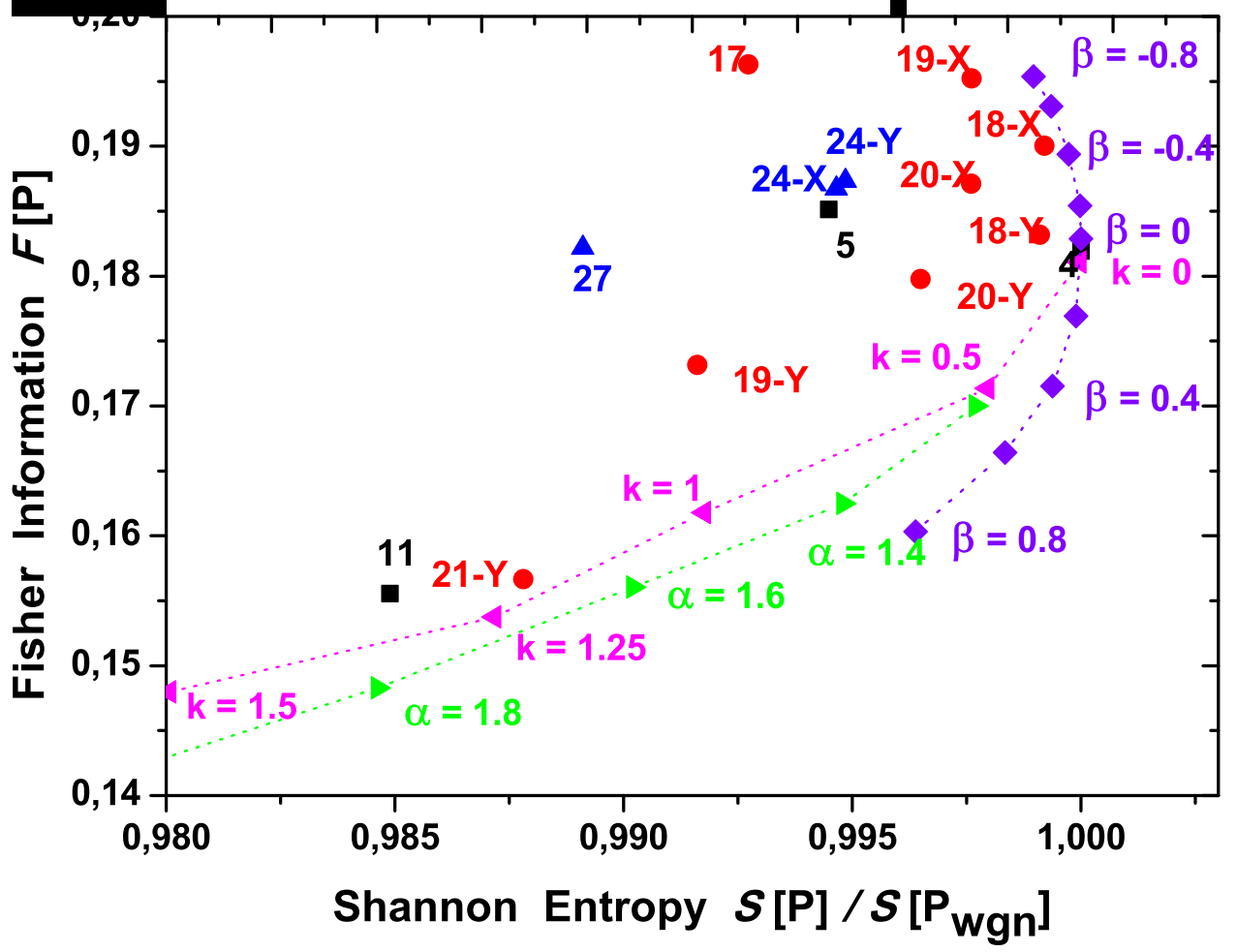


Figure 11. Shannon-Fisher plane, $\mathcal{S} \times \mathcal{F}$ zoom, see Fig. 9